

Running head: EVOLVING FACES

Evolving faces from principal components

Peter J.B. Hancock
Department of Psychology
University of Stirling
FK9 4LA, UK
pjbh1@stir.ac.uk
+44 1786 467675

Abstract

A system that uses an underlying genetic algorithm to evolve faces in response to user selection is described. The descriptions of faces used by the system are derived from a statistical analysis of a set of faces. The faces used for generation are transformed to an average shape by defining locations around each face and morphing. The shape-free images and shape vectors are then separately subjected to principal components analysis. Novel faces are generated by recombining the image components ("eigenfaces") and then morphing their shape according to the principal components of the shape vectors ("eigenshapes"). The prototype system indicates that such statistical analysis of a set of faces can produce plausible, randomly generated photographic images.

Behavior Research Methods, Instruments and Computers 32-2, 327-333, 2000

Introduction

When someone is witness to a crime, a key need is to extract a description of the suspect. Verbal descriptions usually give only a rough idea and attempts to generate a likeness of the target face have traditionally centred on composite methods. The witness is asked to select from a number of sets of eyes, noses, mouths etc. The technology for doing this has progressed from identikit, through photofit¹, to computer-based systems such as EFIT² (Shepherd and Ellis 1992) and the closely related CD-Fit³, growing in sophistication and ease of use. However, the method remains flawed in that humans do not perceive faces as individual facial features: change the eyes and our perception of the nose may change; widen the mouth and the whole face may look narrower. Formal tests of photofit showed that people had difficulty in generating a recognisable likeness even with the target in front of them (Ellis, Shepherd and Davies 1975). The more modern systems allow the shape of features to be altered, but are still based on composites. Since we perceive faces holistically, a more holistic means of face generation may facilitate accurate reconstruction.

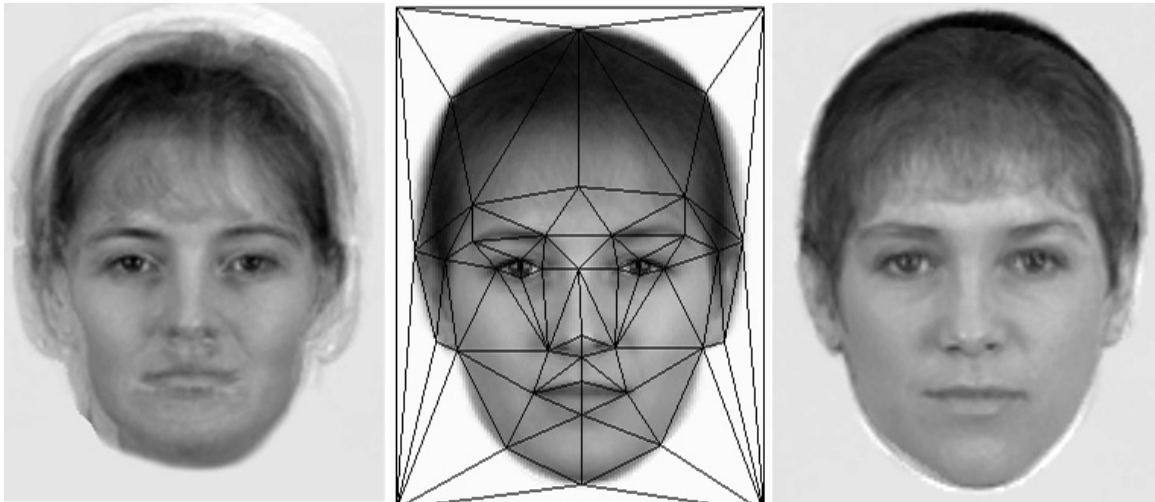
Principal components analysis

An inherently holistic representation for faces is principal component analysis (PCA) (Kirby and Sirovich 1990), which has been used in identification systems (Turk and Pentland, 1991). A number of target face images are aligned, e.g. by scaling and translation, to put the eyes in the same x,y location. They are then analysed to give the eigenvectors, commonly known as eigenfaces. Some examples are shown in [Figure 1](#). Recombination of these eigenfaces in the correct proportions will recover the original faces. These proportions form the characteristic signature of the face. If a new face is analysed using the same eigenface set, it will produce a signature that can be matched to those of the targets, usually by using the simple Euclidean distance. This is the basic methodology of PCA-based face identification.



[Figure 1](#). The first four eigenfaces from this set of 20 faces.

If eigenfaces are recombined in proportions other than those characteristic of the target faces, new images will be created. This is the method underlying the tool described by O'Toole and Thompson (1993), which allows users to vary the proportions of each component in the displayed face. Using simple PCA on unadjusted faces, novel faces generated in this way are likely to be obviously synthetic. This is because only the positions of the eyes were matched prior to PCA; other features will differ in location and the eigenfaces have to accommodate this variation. A random combination is therefore likely to have very blurred features, see [Figure 2a](#).



[Figure 2](#). a) A face generated by random combination of components from originals matched only for location of eyes. b) The average shape-free face, showing the points used to locate features. c) A face generated by random combination of shape-free components.

A method to overcome this was suggested by Craw and Cameron (1991). Prior to PCA, they mark the location of key features around each face and morph them to an average shape, [Figure 2b](#). Now the space for PCA is more linear and random recombination of the eigenfaces produces a much sharper result, see [Figure 2c](#). However, the faces generated all have the same, average shape. This can be rectified by performing PCA on the shape vectors - the x,y locations of all the features. This produces “eigenshapes”, the effects of which can be seen by applying them to the average shape free face. Animations of some sequences from a different set of faces can be seen at <http://www.stir.ac.uk/psychology/Staff/pjbh1/facepca.html>. In that set, the first eigenshape codes the angle of the head - the animation is seen to be nodding. Despite asking people to look straight at the camera, the major source of variation turns out to be head angle. The set of faces used here appears to be rather better controlled in this respect, so the first eigenshape codes overall face size. This is illustrated in [Figure 3](#), which shows deviations either side of the average generated by the first five eigenshapes. Note that the deviations shown are rather larger than are found in the sample population, in order to make clear what is being coded. The second component codes face width. The third component codes the tilt of the head, unfortunately mixed in with something to do with the relative width of the top and bottom of the face. The

likelihood of such multi-modal shape variation reduces with bigger image samples. The fourth component here is also complex, with left-right turning of the head mixed with variations of shape of mouth and nose. The most obvious effect of the fifth is on the position of the eyebrows, though eyes and ears move as well.

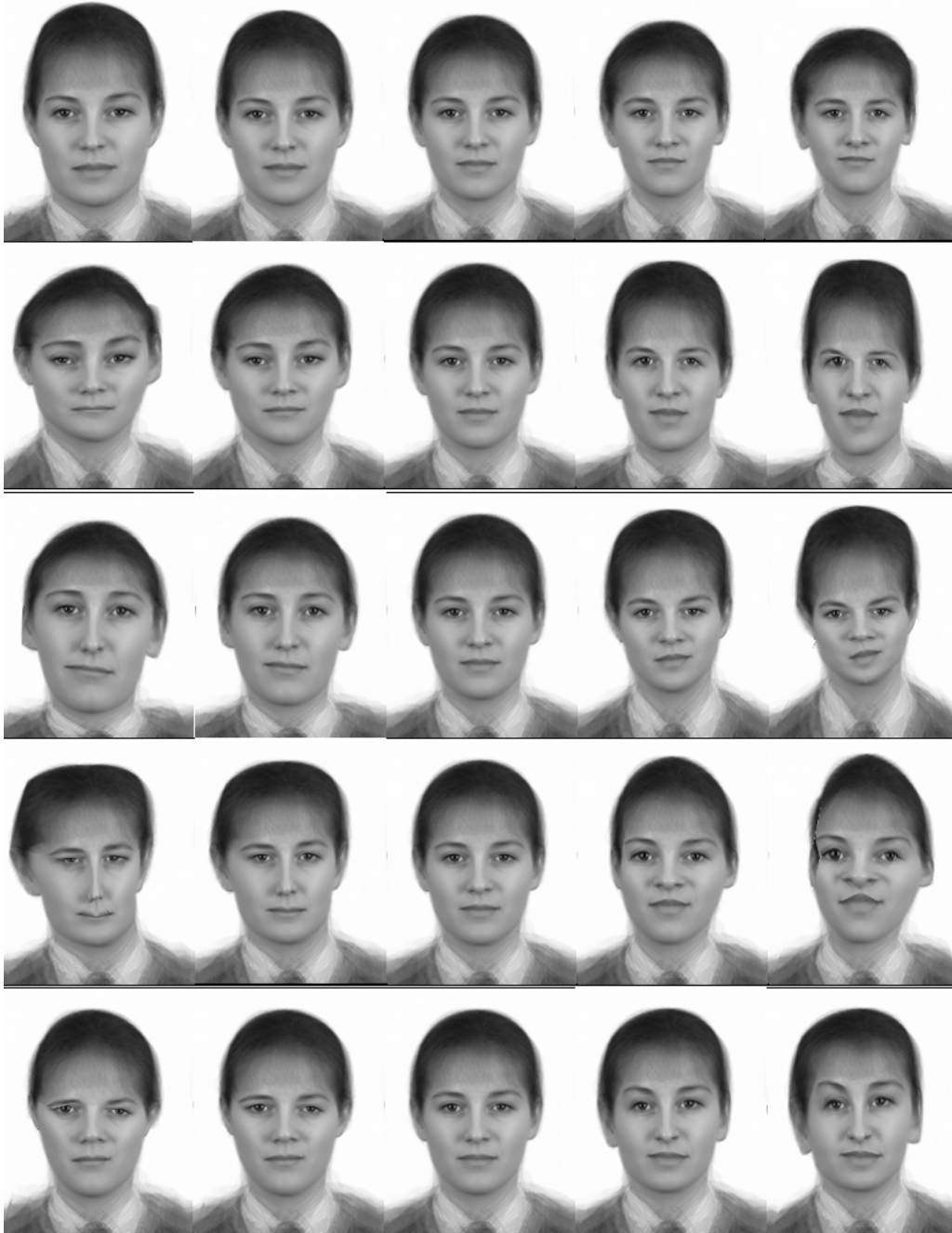


Figure 3. Illustration of the effects of the first five eigenshapes. The central image in each row is the average shape-free face, either side are variations produced by applying the eigenshape to the average face co-ordinates and morphing to the new shape. The amplitude of the variation is large to make the effects clear.

Craw and Cameron (1991) showed that separating out shape information in this way lead to improved identification of new faces. Hancock, Burton and Bruce (1996) showed that it also improves the match with human notions of distinctiveness and memorability for the same faces.

Recreation of one of the original targets requires combining the image components in their appropriate proportions and then morphing back to the original shape by applying the required shape components. Plausible looking novel faces can now be generated by random recombination of the shape-free eigenfaces, followed by random application of the eigenshapes. We now have a means of generating faces holistically rather than by piecing together individual features.

Genetic search

One plausible way to explore this space of faces would be to create an interface with a set of “knobs” that controls the amount of each eigenface and eigenshape in the generated face, similar to the eigenface system described by O'Toole and Thompson (1993). To do so would require considerable computational power, since any change would require the face to be regenerated from scratch. However, it would also be rather bewildering to use, since there might typically be 50 image and 20 shape components.

The slider method is used by the “SpotIt” face generation system (Brunelli and Mich, 1996). They divide a face up into feature regions, such as the nose and mouth, and use PCA to examine the variation within each local region. They then use a set of sliders to control the amount of each eigenfeature mixed into the image. Their system is an interesting hybrid, still essentially using a composite approach, but one that gives smooth variations on each feature, rather than the stepped change of traditional composites.

An alternative method to explore the space is suggested by the evolutionary user selection program described by Dawkins (1986). This program produced a number of 2D branching structures, controlled by a set of parameters. The user would indicate preferences and an underlying genetic algorithm shuffled the selected parameters to produce new "biomorphs" for the next generation. This simple system allowed quite complex designs to be produced, demonstrating the power of selection. Subsequently, the method has been developed to produce more sophisticated computer graphics (e.g. Sims 1991). Caldwell and Johnston (1991) extended the idea to generating face images. Their system presented a set of randomly generated photofit images to a user. These were rated according to similarity with the target face. These ratings were then used by the underlying algorithm to generate a new set of faces, weighted towards those features selected by the user. In this way the user was able to evolve a face that matches the target.

Johnston has continued using such evolutionary design methods for face generation, attempting to produce faces rated as beautiful (Johnston and Franklin, 1993). He also has a web site that allows users to contribute to an online experiment. These still use composites to generate the faces. A number of other online systems have appeared from

time to time, but so far as I am aware no-one else has used PCA in this way to generate the images.

Method

Materials

The source faces need to be photographed under consistent conditions, otherwise PCA will extract components that describe the variation in the conditions, such as lighting, rather than the faces. The source images used here were supplied by the UK Home Office and are of policewomen, who conveniently tend to have short or tied-back hair. 20 faces were used for this trial, avoiding any that had longer hair. The required resolution is not high: the images used here started at Photo-CD resolution of 1536x1034 pixels, including quite a lot of surround, but were edited down to 122x160. The background was set to pure white using a photo-retouching program. Any extraneous wisps of hair were edited out, along with anything below the jaw line.

Shape

Key locations around each face were marked by visual inspection ([Figure 2b](#)). For this work, 38 points were defined, although three of these were defined purely by reference to others. These were the three across the forehead, which were set a specified number of pixels above those locating the ends of the eyebrows and the bridge of the nose. This was done because it was found that visually acceptable morphing seemed to require control points in this region, but there were no obvious landmarks to locate them. The points were chosen to do earlier work investigating the utility of PCA for face identification and there is no reason to think that they are particularly suitable for the purpose described here. The triangles between the points were defined by hand (in a file that holds the list of which points define each triangle) and are in the same pattern for each face.

Once all the faces had been triangulated, the shape files were simply averaged to give the mean position for each point. This average face shape also marks the average position for the faces, which varies since the bounding box for each face was drawn by hand during the editing process. Each face was then morphed to the average shape, using a home-written program developed before commercial programs became available. This works by calculating a new grey-level for each pixel within each triangle in turn, by calculating where in the original image it would come from and using bicubic interpolation to calculate the grey level at that point. The process requires a second or two of CPU time, on a now elderly workstation, per image.

Once all the faces are the same shape, and in the same location in the image, they may simply be averaged. This gives an average shape-free face, shown in [Figure 2b](#). This forms the base face, to which deviations identified by the PCA are added.

Principal Components Analysis

PCA is performed on the shape-free image set. The program delivers the set of eigenfaces in separate files for display, along with the whole set in a large matrix suitable for loading

into the face generation system. It also reports the associated eigenvalues, which are used to control the allowed variation in the generation system.

PCA is also performed on the set of shape vectors. Since these record the position of each face as well as its shape, they are first normalised to bring the point marking the tip of the nose to a constant position. Some of my earlier work on PCA of faces normalised them for the centre of each eye. This had the advantage that it could do some adjustment of size and head angle, but had the unfortunate effect of making all the variations relative to the distance between the eyes, which is perceptually unsettling. The variations produced by the eigenshapes here are illustrated in

Figure 3. Any eigenshapes coding primarily for orientation are excluded from the regeneration process, since such variations only serve to confuse. Here, just the first vector (size) is excluded, since there is no point of reference to judge head size in any case, while later orientation components are mixed in with face shape changes.

Face generation

We now have the basis for generating new faces. A face is produced by adding eigenfaces to the average shape-free face and then distorting it by applying eigenshapes. The code for a particular face is given by the proportions of each eigenface and eigenshape involved in its creation. It is this code - a set of real numbers - which is manipulated by the underlying evolutionary algorithm. Initially, a random set of faces is generated using Gaussian random variables for each parameter. The standard deviation of the Gaussian is proportional to the eigenvalue of the relevant eigenvector, image or shape. In this way the population of parameter values approximates the distribution of the original faces, since the early components carry far more of the variance than later ones. For this prototype, only 10 image and 9 shape components were used (10, minus the one coding for face size).

Figure 4 shows the output of the pilot system on initial start-up, with faces generated entirely at random. The faces produced are mostly plausible as photographs of real people, with only a couple looking somewhat distorted. There remain many issues about the system to be resolved, one being the size of the initial spread of parameter values. Over-large shape variations are the likely cause of distorted looking faces. Some of the faces look quite similar, a function of the limited size of the generating set for this trial. A full generating set might require of the order of 100 faces. The overall range of faces produced here is also quite limited, again a function of the generating set. There are also some visual oddities around the edge of the faces, especially around the hair. Handling hair satisfactorily seems likely to be one of the biggest difficulties in developing a full system.

In this prototype, the user indicates preferences by means of a slider underneath each face, with a scale from 0-10. These evaluations are used by the underlying evolutionary algorithm to select parents for the next generation in simple proportion to the assigned ratings. Baker's algorithm is used to remove undesirable stochastic effects (Baker 1987). Parent strings are chosen in pairs from those selected, to produce pairs of offspring. Uniform crossover is applied, with a probability defined at run time, 0.5 in the examples shown. Note that the parameter values are held as real numbers, so crossover acts only

between these values and not within them, as it would in a bit string genetic algorithm (Holland 1975). Mutation is applied to each variable with a probability of 0.5, set artificially high here to illustrate the effects. Mutations take the form of a Gaussian random variable, with a standard deviation proportional to the eigenvalue of the principle component in question. This produces relatively large variations in the components that carry a lot of the variance and smaller ones in those that normally vary little. The new parameters are used to generate a fresh set of faces, a process that takes a second or two per face in the prototype.



Figure 4. Randomly generated faces produced by the pilot system.

To illustrate the variations generated by the system, Figure 5 shows six faces produced purely by random mutations of the top left face of Figure 4. Figure 6 shows variations produced by crossover alone, acting on the top left and middle left faces of Figure 4. Finally, Figure 7 shows variations produced by the action of both crossover and mutation on the same two original faces.



Figure 5. Variations on the first face of Figure 4 produced by mutation alone.

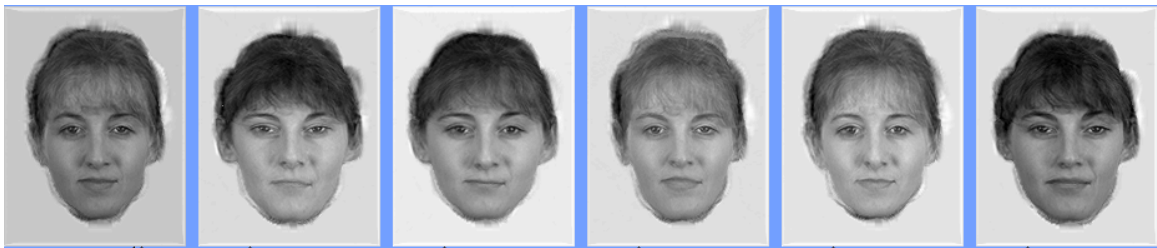


Figure 6. Variations produced by crossover alone acting on the top left and middle left faces from Figure 4.



Figure 7. Variations produced by both crossover and mutation acting on the top left and middle left faces from Figure 4.

Future work

What has been described is very much a prototype system, designed to illustrate the possibility of generating faces in this way. A project is now underway to develop and test a full working version. A number of issues require investigation.

For a start, it isn't clear whether such an evolutionary method will be good from a psychological point of view. We simply don't know whether being presented with a set of faces in this way will help or hinder the user in recall. It may be too confusing, perhaps especially as the system converges, with many close variations. Users may also wish to fix a feature, such as a nose, which would not readily be possible with the system as described here.

Even if the method is helpful, there will be many issues to do with the evolutionary algorithm and the underlying coding to resolve. How many faces should be presented? What are appropriate selection pressures, crossover and mutation rates? Ought users to rate each face, or simply select a few they think most like the target? What is an acceptable image resolution? How varied should the set of faces used for generating the PCA be? It seems likely that an initial narrowing down of the range of targets should be done, limiting age, race and gender. Clearly the system will not produce faces outside the space of those used to generate the coding. However, for mixed-race targets, the PCA approach may be able to produce more accurate images than a more traditional composite system, which may not have the necessary intermediate features available.

The images shown in [Figure 4](#) reveal a number of flaws, especially around the edges of the faces. Hair is clearly problematic, due to its sheer variety of appearances. A simple statistical analysis offered by PCA may not be good enough and some sort of composite system may be needed. Other edge effects may be improved by an alternative set of control points for the face.

One of the aims of this project is to improve our understanding of the way we may represent faces internally. The notion is that the better the match between the space manipulated by the face generation system and that used by our brains, the easier it should be to generate desired faces. This system thus gives an inverted way to test possible representations. Although the shape-free PCA coding has been shown to have significant correlations with human perceptions of faces (Hancock et al 1996), the correlations are not very high and initial playing with the prototype suggests that they may not be the final answer for this form of face generation either. We intend to investigate combinations of global and more local coding methods, perhaps using localised PCA or Independent Components Analysis (Bartlett & Sejnowski, 1997).

An alternative approach would be to attempt to build in some of our knowledge of the way that faces grow. For example, Pittenger and Shaw (1975) showed that a relatively simple mathematical transformation (cardioid strain) gives a good account of the major changes as a head grows from childhood to adult. While our population are, currently, all adult, it may be that growth constraints imply covariations of feature sizes and positions that could be specified mathematically, rather than derived statistically. For example, we might try to model the effects of testosterone, which produces a characteristic pattern of facial development (e.g. enlarging the jaw and brow ridge (Thornhill and Gangestad, 1996)). We suspect that a move to 3D descriptions of faces may be necessary in order to progress these ideas.

Acknowledgements

This project is funded by EPSRC grant no. GR/L88627. Bob Nicholls at the UK Home Office provided the original face images used.

References

- Baker, J. E. (1987). Reducing bias and inefficiency in the selection algorithm. In J.J. Grefenstette (Ed), Proceedings of the Second International Conference on Genetic Algorithms (pp 14-21). Lawrence Earlbaum.
- Bartlett, M. S. & Sejnowski, T. J. (1997) Independent component analysis of face images: A representation for face recognition. Proceedings of the Fourth Annual Joint Symposium on Neural Computation. La Jolla, Ca: Institute for Neural Computation.
- Brunelli, R. & Mich, O. (1996) SpotIt! An interactive Identikit system. Graphical Models and Image Processing, 58(5), 399-404.
- Caldwell, C. & Johnston, V. S. (1991). Tracking a criminal suspect through "face space" with a genetic algorithm. In R.K. Belew & L.B. Booker (Eds), Proceedings of the Fourth International Conference on Genetic Algorithms (pp 416-421). San Mateo, Ca: Morgan Kaufmann.
- Craw, I. & Cameron, P. (1991). Parameterising images for recognition and reconstruction. In P. Mowforth (Ed), Proceedings of the British Machine Vision Conference (pp 367-370). London: Springer-Verlag.
- Dawkins, R. (1986) The Blind Watchmaker. Harlow: Longman.
- Ellis, H., Shepherd, J. & Davies, G. (1975) An investigation into the use of the photo-fit technique for recalling faces. British Journal of Psychology, 66(1), 29-37.
- Hancock, P. J. B., Burton, A. M. and Bruce, V. (1996) Face processing: human perception and principal components analysis. Memory and Cognition, 24, 26-40.
- Holland, J. H. (1975) Adaptation in Natural and Artificial Systems. Ann Arbor, MI: University of Michigan Press.
- Johnston, V. S.. & Franklin, M. (1993) Is beauty in the eye of the beholder? Ethology and Sociobiology, 14(3), 183-199.
- Kirby, M. and Sirovich, L. (1990) Applications of the Karhunen-Loeve procedure for the characterization of human faces. IEEE: Transactions on Pattern Analysis and Machine Intelligence, 12(1), 103-108.
- O'Toole, A. J. & Thompson, J. L. (1993) An X Windows tool for synthesizing face images from eigenvectors. Behavior Research Methods, Instruments and Computers, 25, 41-47.
- Pittenger, J.B. and Shaw, R.E. (1975) Aging faces as viscal elastic events: Implications for a theory of nonrigid shape perception. Journal of Experimental Psychology: Human Perception and Performance 1, 374-382

Shepherd, J. & Ellis, H. (1992) Face recognition and recall using computer-interactive methods with eye-witnesses. In V. Bruce and A. M. Burton (Eds) Processing Images of Faces (pp 129-148). Norwood, NJ: Ablex Publishing Company.

Sims, K. (1991) Artificial evolution for computer graphics. Computer Graphics, 25, 319-328.

Thornhill, R. and Gangestad, S.W. (1996) The evolution of human sexuality. Trends in Ecology and Evolution 11(2), 98-102

Turk, M. & Pentland, A. (1991). Eigenfaces for recognition. Journal of Cognitive Neuroscience, 3, 71-86.

Footnotes

¹ Photofit was devised by Jaques Perry and marketed by John Waddington of Kirkstall, Leeds, UK

² EFIT is marketed by Blaze Imaging Ltd, Pentland House, Livingstone UK. <http://www.blazeimaging.com>

³ CD-fit is marketed by ZEDA ABM, Scion House, Stirling University Innovation Park, Stirling, FK9 4NF. <http://www.zeda-abm.com>